AFRL-VA-WP-TP-2006-311

ON THE COMPUTATION OF THE EGO-MOTION AND DISTANCE TO OBSTACLES FOR A MICRO AIR VEHICLE (PREPRINT)

R.V. Iyer

JANUARY 2006

Approved for public release; distribution is unlimited.

STINFO INTERIM REPORT

AIR VEHICLES DIRECTORATE
AIR FORCE RESEARCH LABORATORY
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7542

NOTICE

PAO Case Number: AFRL/WS 06-0190, 24 Jan 2006.

THIS TECHNICAL REPORT IS APPROVED FOR PUBLICATION.

/s/

Phillip Chandler
Senior Aerospace Engineer
Control Design and Analysis Branch
Air Force Research Laboratory
Air Vehicles Directorate

/s/

Deborah S. Grismer
Chief
Control Design and Analysis Branch
Air Force Research Laboratory
Air Vehicles Directorate

/s/

Brian W. Van Vliet
Chief
Control Sciences Division
Air Force Research Laboratory
Air Vehicles Directorate

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE ***(DD-MM-YY)***	**2. REPORT TYPE**	**3. DATES COVERED** ***(From - To)***
January 2006	Conference Paper Preprint	06/03/2005 – 09/30/2005

4. TITLE AND SUBTITLE ON THE COMPUTATION OF THE EGO-MOTION AND DISTANCE TO OBSTACLES FOR A MICRO AIR VEHICLE (PREPRINT)	**5a. CONTRACT NUMBER** In-house **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER** N/A
6. AUTHOR(S) R.V. Iyer	**5d. PROJECT NUMBER** N/A **5e. TASK NUMBER** N/A **5f. WORK UNIT NUMBER** N/A
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Control Design and Analysis Branch (AFRL/VACA) Control Sciences Division Air Vehicles Directorate Air Force Research Laboratory, Air Force Materiel Command Wright-Patterson AFB, OH 45433-7542	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFRL-VA-WP-TP-2006-311
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Vehicles Directorate Air Force Research Laboratory Air Force Materiel Command Wright-Patterson Air Force Base, OH 45433-7542	**10. SPONSORING/MONITORING AGENCY ACRONYM(S)** AFRL/VACA **11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)** AFRL-VA-WP-TP-2006-311

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

This report contains color.

This work has been submitted to the 2006 IEEE American Control Conference proceedings. This is a work of the U.S. Government and is not subject to copyright protection in the United States.

PAO Case Number: AFRL/WS-06-0190 (cleared January 24, 2006).

14. ABSTRACT

In this paper, we have considered the problem of velocity and range estimation for an UAV using camera and the knowledge of total speed through a GPS device. Together with [5], we have shown that this problem can be solved using a reliability-based motion computation and a optimization problem that is well-posed. If the velocity in the body frame is known, then the problem results in a straight-forward solution. For the more complicated case, when only the total speed is known, we assume that the component of the linear velocity along the axis of the camera is positive. We show that these assumptions form minimal additional information required to solve the problem of ego-motion and range estimation.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Phillip Chandler
Unclassified	Unclassified	Unclassified	SAR	20	**19b. TELEPHONE NUMBER** *(Include Area Code)* (937) 255-8680

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

On the Computation of the Ego-Motion and Distance to Obstacles for a Micro Air Vehicle *

R. V. Iyer

Abstract

In this paper, we have considered the problem of velocity and range estimation for an UAV using camera and the knowledge of total speed through a GPS device. Together with [5], we have shown that this problem can be solved using a reliability-based motion computation and a optimization problem that is well-posed. If the velocity in the body frame is known, then the problem results in a straight-forward solution. For the more complicated case, when only the total speed is known, we assume that the component of the linear velocity along the axis of the camera is positive. We show that these assumptions form minimal additional information required to solve the problem of ego-motion and range estimation.

1 Introduction

In this paper, we have considered the problem of velocity and range estimation for an UAV using vision based techniques. Such a problem is of great importance to Micro Air Vehicles (MAVs) that fly at low enough altitudes so that GPS geo-registration errors can cause them to fly into obstacles. Loss of GPS during flight for short periods of time could also result in loss a MAV. For such applications such as search and classify targets, it is necessary for MAV's to carry an onboard camera that streams video signals to a stationary receiver. The question that naturally arises is whether the video data can help the MAV navigate in the presence of obstacles. Recent work along these lines can be found in [9, 7, 8].

The problem can be broken down naturally into three subproblems:

1. Estimation of motion flow field in the image plane of the camera in an unconstrained environment (that is, the camera is not made to move in a controlled manner);
2. Estimation of the linear and angular velocities of the MAV and the current range of the objects in the visual field; and
3. Navigation to the desired way-points while avoiding the obstacles that might cause the loss of the MAV.

In this paper, we address the second problem listed above. The fundamental question that needs to be answered for this sub-problem is whether the linear and angular velocities of the MAV and the current range of the objects in the visual field of an MAV can be computed correctly, assuming that the first question has been solved correctly. Assuming that component of the linear velocity along the axis of the camera is positive, we show here that subproblem 2 is solvable correctly, provided additional information on the speed of the MAV.

*Final Report for 2005 ASEE/AFOSR SFFP. Will be submitted to IEEE ACC 2006

Conference paper preprint submitted for publication in the 2006 IEEE American Control Conference Proceedings).

This paper is to be studied together with [5] which tackles the first problem of motion estimation using a reliability-based correspondence computation scheme that applies to successive frames of a video stream. Here we define the notion of a set of *nonsingular structure blocks* that is necessary for the solution of the ego-motion problem. Other techniques for the computation of the flow field on the image plane include the optical flow [4, 10, 7] and scale-invariant feature tracking methods (see [6] and references therein). As the optical flow computation can result in wildly inaccurate solutions (see [4] for a discussion), and in light of Theorem 3.2 it seems that a feature tracking method in some sense is necessary. As there can be translation, rotation and scaling of the image from one frame to the next, it is clear that a scale-invariant approach will be more fruitful. The theoretical results of this paper do not depend on which of the specific motion-field computation methods are used, though we use the notation and terminology employed in [5].

Polat and Pachter [9] while considering the problem of INS aiding using a camera came to a similar conclusion to ours of the necessity of knowing the speed of the aircraft to determine range. However, their analysis was essentially one-dimensional, while we study the full three-dimensional case. In [7], there are several assumptions made that are stronger than that required in this paper. In addition to assuming the direction of flight to be along a coordinate axis of a camera-centered coordinate system, they also assume constant altitude and assume a priori knowledge of the probability distribution for the altitude function. They also assume a certain smoothness in the depth function from one frame to the next. This makes the flight as well as the scenery fairly constrained, though it must be mentioned that the authors of [7] are trying to solve a different problem of estimating the height above ground. Here we are interested in finding what minimal additional information is required during *unconstrained* flight for the computation of ego-motion and range.

In Section 2 - 3, we derive the basic equations of camera motion using general notation that is used in basic aerodynamics texts such as Etkin [3]. The advantage of this is that we can work in mixed coordinates for the angular velocity and linear velocity that is very convenient. For the linear velocity we use the body axes of the aircraft, while for the angular velocity we use a camera-centered coordinate frame. We also allow for the camera to be pointed in a general direction on the aircraft, and for the possibility of it being gimballed. We believe that the simplicity of our modeling procedure compares favorably to other approaches [2] is of interest on its own. Section 3 contains the main results of the paper. One key result is Theorem 3.3 that gives necessary and sufficient conditions for the solution of the ego-motion problem.

2 Kinematics of Camera Motion

There are several reference frames employed in air vehicle computations. The main ones are the inertial frame and the body frame which is centered on the center of mass. When a camera is used on a UAV, one introduces an additional frame that is centered on the focus of the camera. If the camera is fixed, then the change of coordinates from the body frame to the camera-centered frame is accomplished by a fixed rotation and translation. This is the case we assume in this note.

Figure 1 shows an inertial frame with origin at the point O_i, an UAV body frame with origin at the point O_b and the camera-centered frame origin at the point O_c. For the inertial and UAV body frames the standard convention for labelling the axes is assumed [3] with X_b pointing through the nose of the aircraft, Y_b pointing out the right wing, and Z_b pointing down. For the camera frame, the standard convention used in the machine vision literature is assumed [4] with Z_c pointing normal to the image plane. Thus if the camera is mounted in front of the aircraft with the image plane parallel to the $Y_b - Z_b$ plane of the aircraft, then X_b and Z_c will be parallel and perhaps collinear.

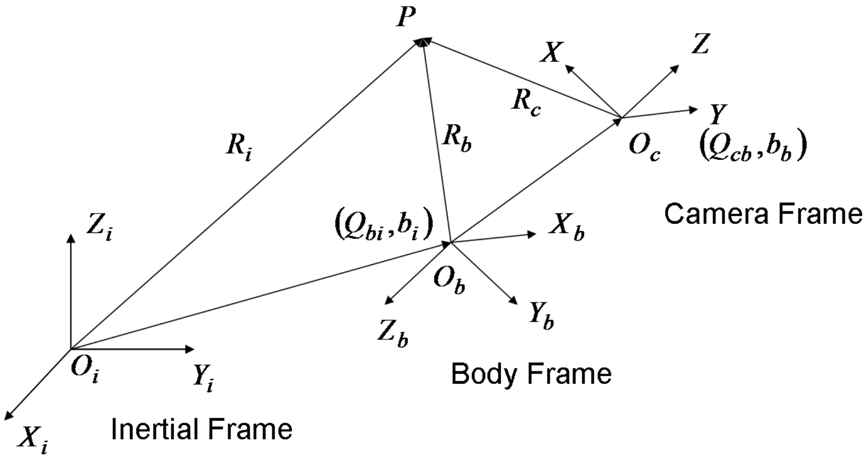


Figure 1: Kinematics of Camera Motion

The point O_b has coordinates denoted by b_i in the inertial frame. The point O_c has coordinates denoted by b_b in the body frame. Suppose that a point P in space has the inertial coordinates R_i, body coordinates R_b and camera-cnetered coordinates R_c, then the relation between these coordinates are given by:

$$R_i = Q_{ib}\, R_b + b_i, \tag{1}$$
$$R_b = Q_{bc}\, R_c + b_b \tag{2}$$

where Q_{ib}, $Q_{bc} \in SO(3)$ are 3×3 matrices satisfying:

$$Q_{ib}^T\, Q_{ib} = I; \text{ and } \det(Q_{ib}) = 1 \tag{3}$$
$$Q_{bc}^T\, Q_{bc} = I; \text{ and } \det(Q_{bc}) = 1 \tag{4}$$

In the following, all variables are assumed to be functions of time unless explicitly stated as constants. Equations (1-2) immediately imply the well-known equations:

$$\dot{Q}_{ib} = Q_{ib}\, \hat{\Omega}_b, \tag{5}$$
$$\dot{Q}_{bc} = Q_{bc}\, \hat{\Omega}_c, \tag{6}$$

where $\hat{\Omega}_b$ and $\hat{\Omega}_c$ are skew-symmetric angular velocity matrices, with the subscript indicating the frame where it is defined. Differentiating Equations (1-2) with the condition that the point P is fixed in the inertial frame, we get:

$$0 = \dot{Q}_{ib}\, R_b + Q_{ib}\, \dot{R}_b + \dot{b}_i \quad \Rightarrow \quad \dot{R}_b = -(Q_{ib}^T\, \dot{Q}_{ib}\, R_b + Q_{ib}^T\, \dot{b}_i) = -(\hat{\Omega}_b\, R_b + V_b) \tag{7}$$

$$\begin{aligned} \dot{R}_b = \dot{Q}_{bc}\, R_c + Q_{bc}\, \dot{R}_c + \dot{b}_b \quad \Rightarrow \quad \dot{R}_c &= -(Q_{bc}^T\, \dot{Q}_{bc}\, R_b + Q_{bc}^T\, \dot{b}_b) - Q_{bc}^T\, (\hat{\Omega}_b\, R_b + V_b) \\ &= -(\hat{\Omega}_c\, R_c + V_c) - Q_{bc}^T\, (\hat{\Omega}_b\, R_b + V_b) \end{aligned} \tag{8}$$

where: $V_b = Q_{ib}^T\, \dot{b}_i$ is the linear velocity of the UAV in the body coordinates, and $V_c = Q_{bc}^T\, \dot{b}_b$ is the linear velocity of the camera in the camera-centered coordinates. If $\Omega_k = [\Omega_{k1}\ \Omega_{k2}\ \Omega_{k3}]^T$ where $k = b$ or c, then:

$$\hat{\Omega}_k = \begin{bmatrix} 0 & -\Omega_{k3} & \Omega_{k2} \\ \Omega_{k3} & 0 & -\Omega_{k1} \\ -\Omega_{k2} & \Omega_{k1} & 0 \end{bmatrix},$$

and it is easy to check that:

$$\hat{\Omega}_k\, R_k = \Omega_k \times R_k. \tag{9}$$

Equations (7-8) model the kinematics of aircraft and camera motion. In special cases, Equation (8) can be simplified:

1. If the camera is fixed on the aircraft, then $\dot{b}_b = 0 \Rightarrow V_c = 0$ and $\dot{Q}_{bc} = 0 \Rightarrow \Omega_c = 0$. This leads to Equation (8) being modified to:

$$\dot{R}_c = -Q_{bc}^T\,(\hat{\Omega}_b\,R_b + V_b).$$

Substituting for R_b from Equation (2), we get:

$$\dot{R}_c = -Q_{bc}^T\,\hat{\Omega}_b\,Q_{bc}\,R_c - Q_{bc}^T\,\hat{\Omega}_b\,b_b - Q_{bc}^T\,V_b = -\widehat{Q_{bc}\,\Omega_b}\,R_c - Q_{bc}^T\,\hat{\Omega}_b\,b_b - Q_{bc}^T\,V_b. \qquad (10)$$

Equation (10) relates the velocity of a point in space in camera-fixed coordinates to linear and angular velocities of the UAV. This is the key equation that will be used in Section 3. Notice that even if the camera is positioned so that the normal to the image plane points along the nose of the UAV, the matrix Q_{bc} is given by:

$$Q_{bc} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}. \qquad (11)$$

2. In case the camera is gimballed and it is possible to control the angular rate of the camera with b_b fixed, we get:

$$\begin{aligned} u &= \Omega_c && (12)\\ \dot{R}_c &= -\hat{u}\,R_c - \widehat{Q_{bc}\,\Omega_b}\,R_c - Q_{bc}^T\,\hat{\Omega}_b\,b_b - Q_{bc}^T\,V_b. && (13) \end{aligned}$$

3 Computation of Motion Parameters from Reliability Indexed Motion Field

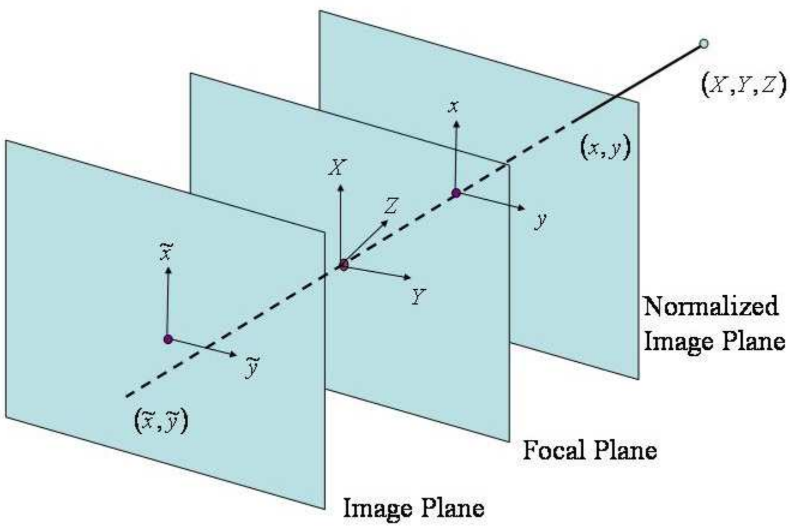


Figure 2: Notation for the Real Motion Field Computation

Figure 2 shows the notation used for the modeling the camera motion. The focal plane contains the focus of the camera and is parallel to the image plane. The origin for the 3 dimensional

coordinate system (X, Y, Z) (henceforth referred to as the camera-centered coordinate system) is at the focus of the camera. As is customary in the machine vision literature, it is assumed that the Z axis is normal to the focal plane. The origin for the coordinate system $(\tilde{x}, \tilde{y})$ on the image plane is located at the point $(0, 0, -f)$ in (X, Y, Z) coordinates, where f is focal length of the camera. The point R with coordinates (X, Y, Z) maps to the point $(\tilde{x}, \tilde{y})$ on the image plane with

$$\tilde{x} = -f\,\frac{X}{Z}, \text{ and } \tilde{y} = -f\,\frac{Y}{Z}.$$

As the image is inverted, we will consider a "normalized" image plane located at $(0, 0, 1)$ and parallel to the focal plane.

Suppose the point $R_c = (X, Y, Z)$ maps to the point (x, y) on the normalized image plane. Then:

$$x = \frac{X}{Z}, \text{ and } y = \frac{Y}{Z}. \tag{14}$$

Equation (9) implies that if the camera moves with linear velocity $V = (V_X, V_Y, V_Z)$ and the angular velocity $\Omega = (\Omega_X, \Omega_Y, \Omega_Z)$, then in the camera-centered coordinate system, the velocity of the point R_c is given by Equation (10): -

$$(\dot{X}, \dot{Y}, \dot{Z}) = \widehat{Q_{bc}\,\Omega_b}\,R_c - Q_{bc}^T\,\hat{\Omega}_b\,b_b - Q_{bc}^T\,V_b. \tag{15}$$

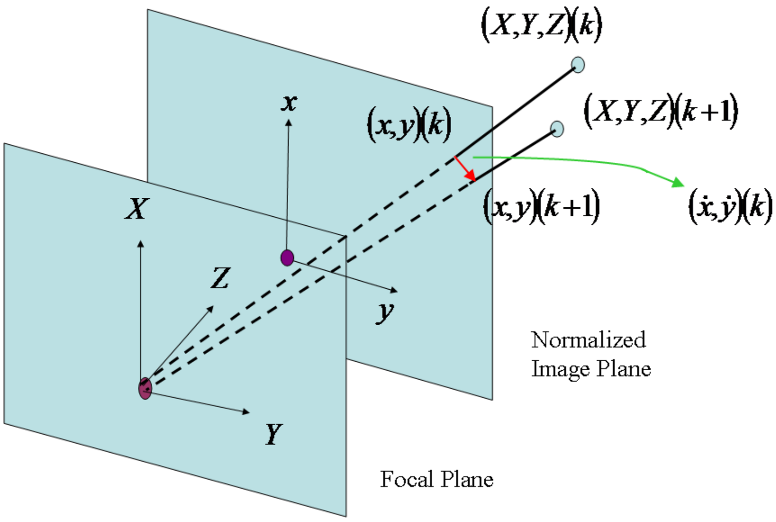


Figure 3: Object motion in camera frame.

Let $Q_{bc} = [Q_{bc1}\ Q_{bc2}\ Q_{bc3}]$ where we have explicitly written the columns of Q_{bc}. Denote the angular velocity Ω_b represented in the camera-centered frame as $\Omega_c = Q_{bc}\,\Omega_b \triangleq [\Omega_X; \Omega_Y\ \Omega_Z]^T$. Then:

$$\begin{aligned}
\dot{x} &= \frac{\dot{X}}{Z} - \frac{X\,\dot{Z}}{Z^2} \\
&= \frac{1}{Z}\,(x\ < Q_{bc3}, V_b > - < Q_{bc1}, V_b >) + \Omega_X\,x\,y - \Omega_Y(1 + x^2) + \Omega_Z\,y \\
&\quad - \frac{1}{Z}\,(< Q_{bc1},\,\hat{\Omega}_b\,b_b > - x\ < Q_{bc3},\,\hat{\Omega}_b\,b_b >) \\
&\approx \frac{1}{Z}\,(x\ < Q_{bc3}, V_b > - < Q_{bc1}, V_b >) + \Omega_X\,x\,y - \Omega_Y(1 + x^2) + \Omega_Z\,y
\end{aligned} \tag{16}$$

$$
\begin{aligned}
\dot{y} &= \frac{\dot{Y}}{Z} - \frac{Y\,\dot{Z}}{Z^2} \\
&= \frac{1}{Z}\,(y\,<Q_{bc3}, V_b> - <Q_{bc2}, V_b>) + \Omega_X(1+y^2) - \Omega_Y\,x\,y - \Omega_Z\,x \\
&\quad -\frac{1}{Z}\,(<Q_{bc1},\,\hat{\Omega}_b\,b_b> - y\,<Q_{bc3},\,\hat{\Omega}_b\,b_b>) \\
&\approx \frac{1}{Z}\,(y\,<Q_{bc3}, V_b> - <Q_{bc2}, V_b>) + \Omega_X(1+y^2) - \Omega_Y\,x\,y - \Omega_Z\,x. \qquad (17)
\end{aligned}
$$

The approximations would be correct if $\frac{b_b}{Z} \approx [0\;0\;0]^T$.

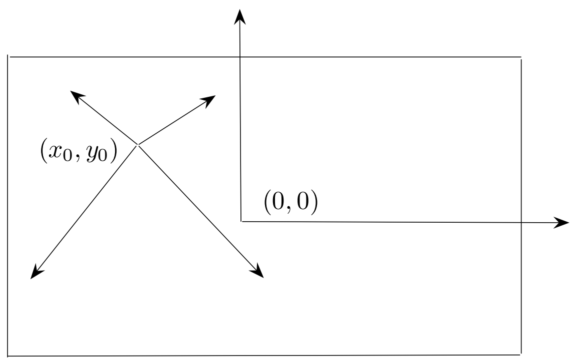
Let $V_b = [V_{b1}\;V_{b2}\;V_{b3}]^T$. For the special orientation of the camera given by (11), we get (assuming $V_{b1} > 0$ which means that the air vehicle has a positive speed along the nose of the air vehicle):

$$
\begin{aligned}
\dot{x} &\approx \frac{\dot{X}}{Z} - \frac{X\,\dot{Z}}{Z^2} \\
&= \frac{V_{b1}}{Z}\left(x + \frac{V_{b3}}{V_{b1}}\right) + \Omega_X\,x\,y - \Omega_Y(1+x^2) + \Omega_Z\,y \qquad (18) \\
\dot{y} &= \frac{\dot{Y}}{Z} - \frac{Y\,\dot{Z}}{Z^2} \\
&\approx \frac{V_{b1}}{Z}\left(y - \frac{V_{b2}}{V_{b1}}\right) + \Omega_X(1+y^2) - \Omega_Y\,x\,y - \Omega_Z\,x. \qquad (19)
\end{aligned}
$$

An erroneous form of these equations appear in [1]. The correct form with a different method of proof appears in [2].

We can make some observations based on these equations.

1. The equations for $\dot{x}$ and $\dot{y}$ are linear in V_b and Ω.
2. If $\Omega = 0$ then $\dot{x} = \frac{1}{Z}\,(x\,<Q_{bc3}, V_b> - <Q_{bc1}, V_b>)$ and $\dot{y} = \frac{1}{Z}\,(y\,<Q_{bc3}, V_b> - <Q_{bc2}, V_b>)$, which means that the point $(x_0, y_0) = \left(\frac{<Q_{bc1},V_b>}{<Q_{bc3},V_b>}, \frac{<Q_{bc2},V_b>}{<Q_{bc3},V_b>}\right)$ is invariant for the instantaneous motion. Note that it is possible for the point (x_0, y_0) to lie outside the boundary of the normalized image plane. For the special case (11), the invariant point turns out to be: $(x_0, y_0) = \left(-\frac{V_{b3}}{V_{b1}}, \frac{V_{b2}}{V_{b1}}\right)$.



Normalized Image Plane

Figure 4: Motion Flow for the case $\Omega = 0$.

3. Another observation for the case $\Omega = 0$ is that at the pixel (x, y) the motion vector $(\dot{x}, \dot{y})$ is directed radially outward from the point (x_0, y_0). The length of this vector is given by:

$$\|(\dot{x}, \dot{y})\| = \frac{< Q_{bc3}, V_b >}{Z} \|(x, y) - (x_0, y_0)\|. \tag{20}$$

   Thus the length of the motion vector is product of the distance of the point (x, y) from (x_0, y_0) and the quantity $\frac{<Q_{bc3}, V_b>}{Z}$.

3.1 Motion Parameter Computation

We will consider two different methods for motion computation. The time interval $[0, T]$ is partitioned into $0 = T_0 < \cdots < T_k < \cdots < T_N = T$. One assumes that linear velocity at time T_{k-1} is known in the body frame and the angular velocity value at time T_{k-1} is computed at time T_k using the images at times T_{k-1} and T_k. In the second method, both linear and angular velocities are computed simultaneously for time T_{k-1}. As discussed in [5], the image is partitioned into structure and non-structure blocks, $\mathbf{B}^m$, $1 \le m \le P$, and the optical motion vector $(\dot{x}^m, \dot{y}^m)$ is computed for the structure blocks located at (x^m, y^m) with say, $m = 1, \cdots, M$.

3.1.1 Method 1

If we had information about the velocity V_b, then the problem becomes very simple and leads to an optimization problem that has an analytic solution. This case is discussed in this subsection. First, we eliminate the pixel-dependent variable Z by suitable transformation. A similar idea (in slightly different contexts) can be seen in [4] and [7].

Combining Equations (18) and (19) and getting rid of the Z, we have

$$D_1(x, y)\dot{x} - D_2(x, y)\dot{y} = C_1(x, y)\Omega_X + C_1(x, y)\Omega_Y + C_3(x, y)\Omega_Z. \tag{21}$$

where

$$\begin{aligned}
D_1(x, y) &= (V_{b1}\, y - V_{b2}), && (22)\\
D_2(x, y) &= (V_{b1}\, x + V_{b3}), && (23)\\
C_1(x, y) &= xy(V_{b1}\, y - V_{b2}) - (1 + y^2)(V_{b1}\, x + V_{b3}), && (24)\\
C_2(x, y) &= -(1 + x^2)(V_{b1}\, y - V_{b2}) + xy(V_{b1}\, x + V_{b3}), && (25)\\
C_3(x, y) &= y(V_{b1}\, y - V_{b2}) + x(V_{b1}\, x + V_{b3}). && (26)
\end{aligned}$$

The unknown variables $(\Omega_X, \Omega_Y, \Omega_Z)$ can be obtained with Least Mean Square Error (LMSE) fitting weighted by the reliability. Let (x^m, y^m) be the pixel coordinate of the center of block $\mathbf{B}^m$. From the motion field analysis, we have obtained the motion vector $(\dot{x}^m, \dot{y}^m)$ for this pixel and the associated reliability measure γ^m. The weight LMS can be written into a matrix form as

$$\mathbf{\Gamma A \Omega} = \mathbf{\Gamma b}, \tag{27}$$

where

$$\mathbf{A} = \begin{bmatrix} C_1(x^1, y^1) & C_2(x^1, y^1) & C_3(x^1, y^1) \\ C_1(x^2, y^2) & C_2(x^2, y^2) & C_3(x^2, y^2) \\ \vdots & \vdots & \vdots \\ C_1(x^M, y^M) & C_2(x^M, y^M) & C_3(x^M, y^M) \end{bmatrix} \tag{28}$$

and

$$\mathbf{b} = \begin{bmatrix} D_1(x^1,y^1)\dot{x}^1 + D_2(x^1,y^1)\dot{y}^1 \\ D_1(x^2,y^2)\dot{x}^2 + D_2(x^2,y^2)\dot{y}^2 \\ \vdots \\ D_1(x^M,y^M)\dot{x}^M + D_2(x^M,y^M)\dot{y}^M \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} \gamma^1 & & & \\ & \gamma^2 & & \\ & & \cdots & \\ & & & \gamma^M \end{bmatrix}. \tag{29}$$

The solution is given by

$$\mathbf{\Omega} = [(\mathbf{\Gamma A})^t(\mathbf{\Gamma A})]^{-1}(\mathbf{\Gamma A})^t(\mathbf{\Gamma b}). \tag{30}$$

3.1.2 Method II

In this method, we show how the linear and angular velocities in the body frame can be computed using Equation 21. Expanding this equation we get the following:

$$P(\Omega, V_b) = a_0 + a_{1x}\,x + a_{1y}\,y + a_{2x}\,x^2 + a_{2xy}\,xy + a_{2y}\,y^2 = 0 \tag{31}$$

where:

$$\begin{aligned}
a_0^m &= -V_{b3}\Omega_X + V_{b2}\Omega_Y + V_{b3}\dot{y}^m + V_{b2}\dot{x}^m; \quad m = 1,\cdots,M & (32)\\
a_{1x}^m &= V_{b3}\Omega_Z - V_{b1}\Omega_X + V_{b1}\dot{y}^m; \quad m = 1,\cdots,M & (33)\\
a_{1y}^m &= V_{b2}\Omega_Z - V_{b1}\Omega_Y - V_{b1}\dot{x}^m; \quad m = 1,\cdots,M & (34)\\
a_{2x} &= V_{b2}\Omega_Y + V_{b1}\Omega_Z & (35)\\
a_{2xy} &= -V_{b2}\Omega_X + V_{b3}\Omega_Y & (36)\\
a_{2y} &= -V_{b3}\Omega_X + V_{b1}\Omega_Z & (37)
\end{aligned}$$

It is not necessary for the coefficients of the polynomial (31) to be zero, because the coefficients of the first three terms change with (x, y).

To solve for the motion parameters (Ω^*, V_b^*), we find the arguments such that:

$$(\Omega^*, V_b^*) = \underset{\substack{(\Omega, V_b) \\ V_{b1} > 0;\ \|V_b\| = V}}{\arg\min}\ J(\Omega, V_b) = \underset{\substack{(\Omega, V_b) \\ V_{b1} > 0;\ \|V_b\| = V}}{\arg\min}\ \sum_{m=1}^{M} |\gamma^m\, P_m(\Omega, V)|^2 \tag{38}$$

Thus we are faced with a minimization problem with constraint $V_{b1} > 0$, and $\|V_b\| = V > 0$ which is the ground speed measured using a Global Positioning System. The idea is to make $|P_m(\Omega, V_b)|$ as small as possible weighted by the reliability of $(\dot{x}^m, \dot{y}^m)$.

Let us now consider the existence of solutions for this optimization problem. The function $P(\Omega, V_b)$ can be written as:

$$P(\Omega, V_b) = V_b^T \left(A(x,y)\,\Omega + B(x,y) \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \right), \tag{39}$$

where

$$A(x,y) = \begin{bmatrix} -x & -y & x^2+y^2 \\ -xy & 1+x^2 & -y \\ -(1+y^2) & xy & x \end{bmatrix}; \quad B(x,y) = \begin{bmatrix} -y & x \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The matrix $A(x, y)$ eigenvalues:

$$0,\ \frac{1}{2}\left(1 + x^2 \pm \sqrt{(1+x^2)^2 - 4\,(1+x^2)y^2 - 4y^4}\right).$$

The right eigenvector corresponding to the 0 eigenvalue is $v_r(x,y) = [x\ y\ 1]^T$, while the left eigenvector is $v_l(x,y) = [-1\ -y\ x]$. It can also be easily checked that $v_l(x,y) \cdot b(x,y,\dot{x},\dot{y}) = 0$.

The physical meaning of $v_l^T(x,y)$ and $v_r(x,y)$ is as follows. Substituting these vectors for V_b and Ω in Equations (18-19) we get $\dot{x} = 0$ and $\dot{y} = 0$. Therefore at every point (x,y) there is an ambiguity in the estimation of the linear velocity in the direction $v_r^T(x,y)$ and the angular velocity in the direction $v_l(x,y)$. However, this ambiguity can be partially resolved by knowledge of $(\dot{x}^m, \dot{y}^m)$ at several points $m = 1, \cdots, M$ where $M \geq 6$. In the equation: $P(\Omega, V_b) = 0$, we still have the issues that if $V_b \neq 0$ is a solution, then any $\alpha\, V_b$ is a solution for $\alpha \neq 0$. This is resolved by the two constraints on the optimization problem, so that we have a unique solution.

We need the following non-singularity condition for the structure blocks.

Definition 3.1 (Non-singularity condition for structure blocks) *The set of structure blocks together with the estimated motion vectors* $\{((x^m, y^m), (\dot{x}^m, \dot{y}^m), \gamma^m);\ m = 1, \cdots, M\}$ *where the reliability indices* $\gamma^m > 0$, *are said to form a non-singular set if for all* $\xi \in \mathbb{R}^3$ *the set of vectors:*

$$\Upsilon \triangleq \left\{ A(x^m, y^m)\, \xi + B(x^m, y^m) \begin{bmatrix} \dot{x}^m \\ \dot{y}^m \end{bmatrix};\ m = 1, \cdots, M \right\}$$

contains at least one non-zero vector.

It is clear that this is a necessary condition for the solution of the problem, because otherwise $J(\Omega, V_b)$ would be zero for some spurious value of angular velocity.

Lemma 3.1 *Assume that the axis of the camera is pointed along the nose of the aircraft, and that the distance of the external objects from the camera is much greater than the distance of the focus of the camera from the center of mass of the aircraft. Then the true velocities* $(\Omega_{b,true}, V_{b,true})$ *satisfy*

$$A(x,y)\, \Omega_{b,true} + B(x,y) \begin{bmatrix} \dot{x}_{true} \\ \dot{y}_{true} \end{bmatrix} \perp V_{b,true}$$

for all (x,y) *in a non-singular set of structure blocks.*

Proof: The first assumption is necessary so that we can use Equations (18 - 19). Let W be the span of set Υ in $\mathbb{R}^3$. Let $V_{b,true} = V_{b,true}^{\perp} + V_{b,true}^{\|}$, where $V_{b,true}^{\perp}$ is the component of $V_{b,true}$ perpendicular to W and $V_{b,true}^{\|}$ is the component along W. We can pick vectors $\{W_i;\ 1 \leq i \leq I;\ 1 \leq I \leq 3\}$ that are orthonormal, and span W. Then:

$$\begin{aligned} V_{b,true}^{\|} &= \sum_{i=1}^{I} \alpha_i\, W_i; \\ A(x^m, y^m)\, \Omega_{b,true} + B(x^m, y^m) \begin{bmatrix} \dot{x}_{true}^m \\ \dot{y}_{true}^m \end{bmatrix} &= \sum_{i=1}^{I} \beta_i^m\, W_i \quad \text{and} \\ J(\Omega_{true}, V_{b,true}) &= \sum_{m=1}^{M} \left(\gamma^m \sum_{i=1}^{I} \alpha_i\, \beta_i^m \right)^2. \end{aligned}$$

Thus $J(\Omega_{true}, V_{b,true}) = 0$ if and only if $\sum_{i=1}^{I} \alpha_i\, \beta_i^m = 0$ for all m, which is true if and only if $V_{b,true}^{\|} = 0$. This proves the claim. □

The above considerations can be summed up in the following theorem:

Theorem 3.1 *Assume that the axis of the camera is pointed along the nose of the aircraft, and that the distance of the external objects from the camera is much greater than the distance of the focus of the camera from the center of mass of the aircraft. Suppose that the true total speed $\|V_{b,true}\|$ is known at any instant of time T_k; $k = 0, \cdots, N$. Suppose that the component of the inertial velocity along the X axis of the camera in Figure 2 is positive, that is, $V_{b1,true} > 0$. Furthermore, suppose that there is no error in the reliability-based motion analysis with the number of structure blocks $M \geq 6$, and that the set of structure blocks form a non-singular set. Denote the true velocities of the air vehicle by $(\Omega_{true}, V_{b,true})$. Then, there exists a unique solution (Ω^*, V_b^*) to the optimization problem (38), and this solution coincides with the true solution $(\Omega_{true}, V_{b,true})$.*

Proof: First observe that the cost function in (38) is quadratic as a function of (Ω, V_b) and that $J(\Omega_{true}, V_{b,true}) = 0$ for the following reason. We substitute $(\Omega_{true}, V_{b,true})$ into Equations (18 - 19) and then substitute the resulting $(\dot{x}_{true}, \dot{y}_{true})$ into (39), which yields $P(\Omega_{true}, V_{b,true}) = 0$. This means that in the absence of noise, the algorithm will converge to a point in the equivalence class $\{(\Omega, V_b) \mid J(\Omega, V_b) = 0\}$. We need to show that this equivalence class consists of only one point $(\Omega_{true}, V_{b,true})$.

The reason for $M \geq 6$ is that there are 6 parameters to be estimated and so we need at least 6 equations for the structure blocks. Another preliminary observation is that the set of points $\{(\Omega, 0);\ \Omega \in \mathbb{R}^3\}$ lead to $P(\Omega, 0) = 0$. However, these points are eliminated by the constraint $\|V_b\| = V > 0$.

If $(\Omega_{true}, V_{b,true})$ is the true solution, and the result of the reliability-based motion estimation (see [5]) is error-free (that is, $(\dot{x}^m, \dot{y}^m)$; $m = 1, \cdots, M$ exactly satisfies Equations (18 - 19)), then we will show that the result of the optimization is $(\Omega^*, V_b^*) = (\Omega_{true}, V_{b,true})$. By rewriting (39) we get:

$$P(\Omega, V_b) = \frac{1}{Z}\, V_b \cdot (v_l \times V_{b,true} + A(x, y)\,(\Omega - \Omega_{true})). \tag{40}$$

Clearly, if $(\Omega, V_b) = (\Omega_{true}, V_{b,true})$ then $P(\Omega, V_b) = 0$. Now suppose $P(\Omega, V_b) = 0$. If $\Omega \neq \Omega_{true}$ then the second term inside the parentheses in (40) is non-zero for a generic point (x, y). It is also a quadratic function of (x, y) by the definition of $A(x, y)$. The first term inside the parentheses is a linear function of (x, y) for a given vector $V_{b,true}$. Hence for a generic point (x, y) the term inside the parentheses is not zero and is a quadratic function of (x, y). As V_b is a constant, $P(\Omega, V_b)$ cannot be zero for a generic point (x, y), which implies that our assumption of $\Omega \neq \Omega_{true}$ is false.

Next suppose that $\Omega = \Omega_{true}$. Then we have:

$$P(\Omega, V_b) = \frac{1}{Z}\, V_b \cdot v_l \times V_{b,true},$$

which is zero for any generic point (x, y) if and only if $V_b = \alpha\, V_{b,true}$, where $\alpha \in \mathbb{R}$. Due to the constraint $\|V_b\| = \|V_{b,true}\|$, we must have $V_b = \pm V_{b,true}$. Now the second constraint $V_{b,1} > 0$ combined with the given condition $V_{b,1true} > 0$ implies that $V_b = V_{b,true}$. □

3.1.3 Effect on Noise on Velocity Estimation for Method II

Next, let us examine the effect of noise on the computation of (Ω, V_b). As observed in [5] there are several sources of noise that affect the computation of $(\dot{x}, \dot{y})$. As the reliability γ^m for structure block m is a measure of the noise in the computed value of $(\dot{x}, \dot{y})$, we consider the reliability analysis to result in random variables (u, v) with mean $(\dot{x}_{true}, \dot{y}_{true})$ and standard deviation $\sigma^m I$. For each m, as $0 < \gamma^m \leq 1$, we consider σ^m to be a continuously differentiable, monotone decreasing function defined on $(0, 1]$ with

1. $\lim_{\gamma \to 0} \sigma^m(\gamma) = \infty$;

2. $\sigma^m(1) = 0$ and

3. $\lim_{\gamma \to 0} \frac{\frac{d\sigma^m(\gamma)}{d\gamma}}{(\sigma^m(\gamma))^2} = \frac{-1}{C}$ for some $C > 0$.

The reason for the last assumption will become clear shortly. For example, $\sigma^m(\gamma) = \tan\left(\frac{\pi}{2}(1-\gamma^m)\right)$ satisfies the requirements with $C = \frac{2}{\pi}$. In practice, the function σ^m could be defined through empirical observations.

For the m-th structure block, denote $(u^m, v^m) = (\dot{x}^m_{true}, \dot{y}^m_{true}) + (\varepsilon^m_x, \varepsilon^m_y)$. Thus $(\varepsilon^m_x, \varepsilon^m_y)$ is a vector random variable with mean $(0,0)$ and standard deviation $\sigma^m I$. Then we have the following theorem:

Theorem 3.2 *Assume that the axis of the camera is pointed along the nose of the aircraft, and that the distance of the external objects from the camera is much greater than the distance of the focus of the camera from the center of mass of the aircraft. Suppose that $V_{b1,true} > 0$. Further suppose that the result of the reliability analysis [5] at time instant T_k is a set of vector random variables (u^m, v^m) with mean $(\dot{x}^m_{true}, \dot{y}^m_{true})$ and standard deviation $\sigma^m I$, where σ^m is related to the reliability γ^m according to conditions 1 - 3 given above. Denote $P_m \equiv P(\Omega, V_b)\Big|_{(x^m, y^m, u^m, v^m)}$. Then for any $\gamma^m \in (0,1]$:*

$$\min_{(\Omega, V_b)} E[\gamma^m P_m]^2 \le \left(\bar{C}\, V_{b1,true}\right)^2 \|(x^m, y^m) - (x_0, y_0)\|^2 \tag{41}$$

where $\bar{C} = \max_{x \in [0,1]} x\sigma^m(x)$ and $(x_0, y_0) = \left(-\frac{V_{b3,true}}{V_{b1,true}}, \frac{V_{b2,true}}{V_{b1,true}}\right)$.

Proof: The theorem follows from straight forward computations. We compute P_m^2 to be:

$$\begin{aligned} P_m^2 &= \left\{V_b^T \left(A(x,y)\Omega + B(x,y) \begin{bmatrix} \dot{x}^m_{true} \\ \dot{y}^m_{true} \end{bmatrix}\right)\right\}^2 + V_b^T B(x,y) \begin{bmatrix} \varepsilon^m_x \\ \varepsilon^m_y \end{bmatrix} \begin{bmatrix} \varepsilon^m_x & \varepsilon^m_y \end{bmatrix} B^T(x,y) V_b \\ &\quad + V_b^T \left(A(x,y)\Omega + B(x,y) \begin{bmatrix} \dot{x}^m_{true} \\ \dot{y}^m_{true} \end{bmatrix}\right) V_b^T B(x,y) \begin{bmatrix} \varepsilon^m_x \\ \varepsilon^m_y \end{bmatrix}. \end{aligned} \tag{42}$$

From this, it is clear that:

$$E[P_m^2] = \left\{V_b^T \left(A(x,y)\Omega + B(x,y) \begin{bmatrix} \dot{x}^m_{true} \\ \dot{y}^m_{true} \end{bmatrix}\right)\right\}^2 + \sigma_m^2\, V_b^T B(x,y) B^T(x,y) V_b. \tag{43}$$

Hence for any $\gamma^m \in (0,1]$:

$$\min_{(\Omega, V_b)} E[\gamma^m P_m]^2 \le E[\gamma^m P_m(\Omega_{true}, V_{b,true})]^2 = (\gamma^m)^2 \left(0 + \sigma_m^2\, V_{b,true}^T B(x,y) B^T(x,y) V_{b,true}\right).$$

We now find an upper bound for the product $\gamma^m \sigma_m(\gamma^m)$. The function $g(x) = x\,\sigma^m(x)$ is defined on $(0,1]$ by the definition of $\sigma^m(\cdot)$. However, it can be seen that $\lim_{x \to 0} g(x)$ is well defined:

$$\lim_{x \to 0} g(x) = \lim_{x \to 0} \frac{x}{\frac{1}{\sigma^m(x)}} = \lim_{x \to 0} \frac{1}{\frac{-(\sigma^m(x))'}{(\sigma^m(x))^2}} \quad \text{(by L'Hospital's Rule)} = C.$$

Hence, by setting $g(0) = \lim_{x \to 0} g(x)$, we have a well defined continuous function defined on $[0,1]$, that has a maximum which we denote by $\bar{C}$. It can be easily checked that:

$$V_{b,true}^T B(x,y) B^T(x,y) V_{b,true} = (V_{b1,true}\, y^m - V_{b2,true})^2 + (V_{b1,true}\, x^m + V_{b3,true})^2.$$

Therefore for any $\gamma^m \in (0,1]$:

$$\min_{(\Omega,V_b)} E[\gamma^m P_m]^2 \leq \bar{C}^2 \left((V_{b1,true}\, y^m - V_{b2,true})^2 + (V_{b1,true}\, x^m + V_{b3,true})^2\right).$$

This equation leads to the claim. □

The interesting aspect of this theorem is that: if $(x^m, y^m) = (x_0, y_0)$ then $\min_{(\Omega,V_b)} E[\gamma^m P_m]^2 = 0$ no matter what γ^m is! A conclusion is that points near the invariant point for linear motion (x_0, y_0) (which is of course unknown a priori for Method II), are more reliable in terms of the computation of the angular and linear velocities even in the presence of noise! The angular velocity does not matter in this computation. Thus, we have motivated the importance of knowing the linear velocity V_b that was the content of Method I. Notice also the term in the right hand side of (41), and the term in the right hand side of (20)!

To sum up Method II:

Theorem 3.3 *Assume that the axis of the camera is pointed along the nose of the aircraft, and that the distance of the external objects from the camera is much greater than the distance of the focus of the camera from the center of mass of the aircraft. Suppose It is known that $V_{b1} > 0$ - that is, the component of the velocity in the body frame along the nose of the aircraft is positive, and the set of structure blocks form a nonsingular set with $M \geq 6$.*

Then the linear and angular velocities (Ω, V_b) computed by (38) coincides with $(\Omega_{true}, V_{b,true})$ if and only if (i) the total V is known at any instant of time T_k; $k = 0, \cdots, N$. (ii) the vectors $(\dot{x}^m, \dot{y}^m)$ for the structure blocks (x^m, y^m) are estimated correctly and coincide with the true values $(\dot{x}^m_{true}, \dot{y}^m_{true})$.

Proof: The if part of the claim follows from Theorem 3.1. The only if part follows from the proof of Theorem 3.2, where it can be seen by Equations (42-43) that unless $\sigma^m = 0$, the solution of the optimization problem will not coincide with $(\Omega_{true}, V_{b,true})$ in general. □

3.2 Range Estimation

Once the camera motion (Ω, V_b) is computed through either of the Methods I or II, we can determine the range (or depth) Z for each block in the scene. As discussed earlier and detailed in [5], the image is partitioned into blocks, $\mathbf{B}^n$, $1 \leq n \leq P$. If the $(\dot{x}^m, \dot{y}^m);\, ; m = 1, \cdots, M$ are the motion vectors computed for the structure block (x^n, y^n) using the reliability based estimation scheme [5], then the range Z^m for these blocks can be determined by least mean squared error estimation:

$$Z^m = \arg\min_{Z} [\dot{x}^m - f(x^n, y^n, Z)]^2 + [\dot{y}^m - g(x^n, y^n, Z)]^2; \quad m = 1 \cdots, M, \tag{44}$$

where:

$$\begin{aligned} f(x,y,Z) &= \frac{V_{b1}}{Z}(x + \frac{V_{b3}}{V_{b1}}) + \Omega_X xy - \Omega_Y(1+x^2) + \Omega_Z y, & (45)\\ g(x,y,Z) &= \frac{V_{b1}}{Z}(y - \frac{V_{b2}}{V_{b1}}) + \Omega_X(1+y^2) - \Omega_Y xy - \Omega_Z x. & (46) \end{aligned}$$

Let $\Lambda^n = \{(\dot{x}^n_j, \dot{y}^n_j) | 1 \leq j \leq L^n\}$ be the top candidate motion vectors for the n-th *non-structure motion block.* Recall that the non-structure blocks are not used in the computation of (Ω, V_b) and hence we may have multiple vectors for a non-structure block. If the block that corresponds to an object in the scene is stationary, the true motion vector must satisfy Eqs. (18-19) whose right hand

sides are defined in (45-46). Observe that the functions $f(x, y, \cdot)$ and $g(x, y, \cdot)$ are affine functions of $\frac{1}{Z}$ for each x and y. For each candidate motion vector $(\dot{x}_j^n, \dot{y}_j^n)$ for the non-structure block (x^n, y^n), we can compute the corresponding range by orthogonal projection (see Figure 5):

$$Z_j^n = \arg\min_Z [\dot{x}_j^n - f(x^n, y^n, Z)]^2 + [\dot{y}_j^n - g(x^n, y^n, Z)]^2. \tag{47}$$

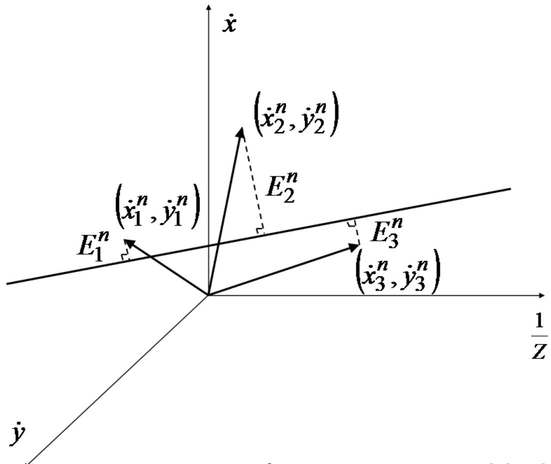


Figure 5: Range Estimation for non-structure blocks

The corresponding fitting error is denoted by

$$E_j^n = [\dot{x}_j^n - f(x^n, y^n, Z_j^n)]^2 + [\dot{y}_j^n - g(x^n, y^n, Z_j^n)]^2. \tag{48}$$

We choose the motion vector in the collection Λ^n to be the one with the least fitting error:

$$j^* = \arg\min_{j=1,\cdots,L^n} E_j^n. \tag{49}$$

The range of the block is given by Z^{j^*}, and the associated motion vector is $(\dot{x}_{j^*}^n, \dot{y}_{j^*}^n)$.

4 Conclusion

In this report, we have considered the problem of velocity and range estimation for an UAV using camera and the knowledge of the total speed of the UAV. Together with [5], we have shown that the ego-motion problem can be solved using a reliability-based motion computation, followed the solution of a well-posed optimization problem. Once the velocities have been found, the range of the objects can be computed easily.

5 Acknowledgement

The author gratefully acknowledges a ASEE/AFOSR Summer Faculty Fellowship during the summer of 2005. This work was performed at Air Force Research Laboratory, Air Vehicles Directorate (AFRL/VACA), Wright-Patterson Air Force Base (WPAFB), OH, together with Dr. Zhihai He, University of Missouri, Columbia, MO, and Mr. Phillip Chandler, UAV Team Lead, AFRL/VACA. We also would like to acknowledge discussions with Prof. Meir Pachter of Air Force Institute of Technology, WPAFB, OH, and Prof. David Chichka of George Washington University during the development of this manuscript.

References

[1] J.K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images - a review. In *Proceedings of the IEEE*, volume 76, pages 917–935, August 1988.

[2] Z. Duric and A. Rosenfeld. Shooting a smooth video with a shaky camera. *Machine Vision and Applications*, 13:303–313, 2003.

[3] Bernard Etkin. *Dynamics of Atmospheric Flight.* John Wiley & Sons, New York, 1972.

[4] Oliver Faugeras. *Three-Dimensional Computer Vision - A Geometric Viewpoint.* MIT Press, Cambridge, Massachusetts, 1993.

[5] Z. He. Vision-based UAV flight control and obstacle avoidance. Summer Faculty Fellowship Final Report, Air Force Research Laboratory, Air Vehicles Directorate, WPAFB, OH 45433, 2005.

[6] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, November 2004.

[7] P. Merrell, D. Lee, and R. Beard. Statistical analysis of multiple optical flow values for estimation of unmanned aerial vehicle height above ground. In D. P. Casasent, E. L. Hall, and J. R editors, *Proceedings of SPIE: Intelligent Robots and Computer Vision XXII: Algorithms, Techniques and Active Vision.*

[8] P. Merrell, D. Lee, and R. Beard. Obstacle avoidance for unmanned air vehicles using optical flow probability distributions. In D. W. Gage, editor, *Proceedings of SPIE: Mobile Robots XVII*, volume 5609, pages 13 – 22, Bellingham, WA, 2004.

[9] M. Polat and M. Pachter. INS aiding by tracking an unknown ground object: - theory. In *Proceedings of the IEEE Automatic Control Conference*, June 4-6 2003. Denver, CO.

[10] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger. Probability distributions of optical flow. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 310–315, Maui, HI, June 1991.